

**SNPing the Cost and Proportional–bias error  
for Genetic Identification of Chinook Salmon  
Life Histories from Upstream Rivers to the  
San Francisco Bay Estuary**

**Michael A Banks**

## **Public Comments**

No public comments were received for this proposal.

# Technical Synthesis Panel Review

## Proposal Title

#0240: SNPping the Cost and Proportional–bias error for Genetic Identification of Chinook Salmon Life Histories from Upstream Rivers to the San Francisco Bay Estuary

Final Panel Rating
inadequate

## Technical Synthesis Panel (Primary) Review

### TSP Primary Reviewer's Evaluation Summary And Rating:

This study seeks to apply a new genetic marker type (single nucleotide polymorphism (SNPs)) towards the task of reducing the cost and errors previously associated with microsatellites to identify Chinook runs in the Central Valley. Effective discrimination among these runs will allow for targeted management, protection and recovery efforts.

### Additional Comments:

The reviews ranged from poor to very good. The reviewers were torn in that the idea was good but very poorly communicated. The proposed science was seen as potentially valuable, but a poor and unclear writing style was seen as a huge hindrance to understanding the study objectives and methods. Overall, the proposal was too unclear to warrant funding.

This study seeks to apply a new genetic marker type (single nucleotide polymorphism (SNPs)) towards the task of reducing the cost and errors previously associated with microsatellites to identify Chinook runs in the Central Valley. Effective discrimination among these runs will allow for targeted management, protection and recovery efforts.

#0240: SNPping the Cost and Proportional–bias error for Genetic Identification...

## Technical Synthesis Panel (Discussion) Review

### TSP Observations, Findings And Recommendations:

SNPing the cost and proportional-bias error for genetic identification of Chinook salmon life histories from upstream rivers to the San Francisco Bay Estuary

The primary and secondary reviewers and panel all ranked this proposal as inadequate. While introducing an interesting and potentially new approach for studying fish life histories and distinguishing salmon subpopulations, the proposal was very insufficiently designed and very poorly communicated.

The panel recognized the inconsistencies in the proposal may have resulted from a mistake in file submission. Final Ranking: Inadequate

# Technical Review #1

proposal title: SNPing the Cost and Proportional–bias error for Genetic Identification of Chinook Salmon Life Histories from Upstream Rivers to the San Francisco Bay Estuary

## Review Form

### Goals

Are the goals, objectives and hypotheses clearly stated and internally consistent? Is the idea timely and important?

Comments	The goals and objectives--to develop a cheaper and potentially more reliable genetic marker for Chinook salmon phenotypes--are reasonably clear, whereas the hypotheses are not well developed. The idea, as I perceive the proposal, is interesting. However, this was unquestionably the most poorly written proposal I have yet reviewed, and the scientific justification of the proposed research was, in my view, muddled.
Rating	fair

### Justification

Is the study justified relative to existing knowledge? Is a conceptual model clearly stated in the proposal and does it explain the underlying basis for the proposed work? Is the selection of research, pilot or demonstration project, or a full–scale implementation project justified?

Comments	In essence, the authors propose to discover the genetic basis of run timing in Chinook salmon. I'm not convinced, however, that this is their stated objective, nor am I convinced their methods will achieve this objective, although it's an interesting approach and worth trying. The authors are seeking a genetic signature for each space-time-phenotype (e.g., Feather River-early fall, Clear Creek-spring). This
----------	--

## Technical Review #1

	<p>may be achievable, but not with 100% reliability if environmental factors influence this phenotype to any degree. An ID success rate of greater than 95%, which appears to be the desired outcome, requires an incredibly high degree of genetic penetration, i.e., a near-perfect match between genotype and phenotype. It is unclear whether this is likely to be the case.</p> <p>The conceptual model is not clearly stated, nor is the desired outcome of &gt;95% ID success well justified. However, if SNP analysis can do as well or better than microsatellites at identifying phenotypes, and if it is cheaper, then the SNP approach is justified and therefore worth exploring.</p>
<b>Rating</b>	good

## Approach

Is the approach well designed and appropriate for meeting the objectives of the project? Is the approach feasible? Are results likely to add to the base of knowledge? Is the project likely to generate novel information, methodology, or approaches? Will the information ultimately be useful to decision makers?

<b>Comments</b>	The approach is well designed, and the comparison of microsatellites with SNP markers at timing loci, to disentangle the effects of familial relationships and alleles related to timing, is very interesting.
<b>Rating</b>	very good

## Feasibility

Is the approach fully documented and technically feasible? What is the likelihood of success? Is the scale of the project consistent with the objectives and within the grasp of authors?

## Technical Review #1

<b>Comments</b>	The approach is documented and appears feasible. The methods section appeared well researched, if poorly written.
<b>Rating</b>	good

## Monitoring

If applicable, is monitoring appropriately designed (pre–post comparisons; treatment–control comparisons)? Are there plans to interpret monitoring data or otherwise develop information?

<b>Comments</b>	The research will span multiple years of spawning behavior, to determine the long-term reliability of particular markers in identifying phenotypes.
<b>Rating</b>	very good

## Products

Are products of value likely from the project? Are contributions to larger data management systems relevant and considered? Are interpretive (or interpretable) outcomes likely from the project?

<b>Comments</b>	If the authors can find an equally (or more) reliable and less expensive genetic signature, and perhaps ultimately the genetic basis, of timing behavior in these fish, that will be well worth the funding. Management organizations will benefit by the ability to target particular genotypes for conservation.
<b>Rating</b>	very good

## Additional Comments

<b>Comments</b>	Again, this was the most poorly written proposal I have yet reviewed. It was clearly not proof-read. Run-on sentences, grammatical errors, typographical errors, sloppy
-----------------	---

#0240: SNPing the Cost and Proportional–bias error for Genetic Identification...

## Technical Review #1

	organization, and uninformative figure legends made the proposal very difficult to decipher.
--	--

### Capabilities

What is the track record of authors in terms of past performance? Is the project team qualified to efficiently and effectively implement the proposed project? Do they have available the infrastructure and other aspects of support necessary to accomplish the project?

Comments	The authors have a strong record of funding in this area of research. The quality of the writing diminished their credibility, but I must assume, based on past funding and research success, that this was a fluke.
Rating	very good

### Budget

Is the budget reasonable and adequate for the work proposed?

Comments	Yes, the proposed project spans multiple years and will require a long-term commitment.
Rating	very good

### Overall

Provide a brief explanation of your summary rating.

Comments	Finding a cheaper way to genetically identify Central Valley Chinook salmon phenotypes is worth exploring, and using SNPs in timing loci to do so provides a bonus opportunity to discover the genetic cause of phenotypic variation. I cannot give a high rating, however, because the writing of the proposal did not convince me that the authors have a strong handle on their objectives.
Rating	good

# Technical Review #2

proposal title: SNPing the Cost and Proportional–bias error for Genetic Identification of Chinook Salmon Life Histories from Upstream Rivers to the San Francisco Bay Estuary

## Review Form

### Goals

Are the goals, objectives and hypotheses clearly stated and internally consistent? Is the idea timely and important?

Comments	Goals and objectives of the project are described somewhat poorly. The proposal is disorganized and difficult to read. No hypotheses are formulated, but as the main objectives are choosing useful genetic markers and using these to discriminate individual chinook salmon from different runs, it is not clear what biologically meaningful hypotheses would be. Still, setting up null hypotheses for statistical testing could be useful for framing the objectives and methods.
Rating	fair

### Justification

Is the study justified relative to existing knowledge? Is a conceptual model clearly stated in the proposal and does it explain the underlying basis for the proposed work? Is the selection of research, pilot or demonstration project, or a full–scale implementation project justified?

Comments	The study follows work that has been done in the primary author's lab and in other labs in developing genetic markers with enough variability to allow reasonably accurate identification of individual fish from different runs. One of the proposed ideas, to examine single nucleotide polymorphisms (SNPs) linked to clock genes that might actually be involved in phenotypic divergence between the runs, seems like a
----------	--

## Technical Review #2

	fascinating and promising avenue of research. A conceptual model for the proposed research is not well described, and presentation of background information is neither well organized nor clear. As the proposed tasks are to be performed sequentially, with each task building upon the previous, full-scale implementation seems warranted.
Rating	fair

## Approach

Is the approach well designed and appropriate for meeting the objectives of the project? Is the approach feasible? Are results likely to add to the base of knowledge? Is the project likely to generate novel information, methodology, or approaches? Will the information ultimately be useful to decision makers?

Comments	It is difficult to evaluate the feasibility and appropriateness of the approach because it is not well described in the proposal. If indeed they are able to identify markers that will reduce the cost and increase the accuracy of run assignment of individuals, this would be new information that assumably would be useful to managers. However, there is no justification for this in the proposal - how important is accurate run identification to management at this point? They mention that in the Feather River, differences in run timing have collapsed to a near-continuous arrival of individuals from early September through October. What is the biological significance of this, in terms of recruitment (and thus management)? More justification for the study is needed.
Rating	fair

## Technical Review #2

### Feasibility

Is the approach fully documented and technically feasible? What is the likelihood of success?  
Is the scale of the project consistent with the objectives and within the grasp of authors?

<b>Comments</b>	The approach is not well documented, making feasibility difficult to assess. Methods are not described in sufficient detail. Background information is described poorly.
<b>Rating</b>	poor

### Monitoring

If applicable, is monitoring appropriately designed (pre–post comparisons; treatment–control comparisons)? Are there plans to interpret monitoring data or otherwise develop information?

<b>Comments</b>	Not applicable.
<b>Rating</b>	not applicable

### Products

Are products of value likely from the project? Are contributions to larger data management systems relevant and considered? Are interpretive (or interpretable) outcomes likely from the project?

<b>Comments</b>	The main product would be a suite of loci to be used for discriminating individuals from different runs. As mentioned above, the proposal does not describe the importance of this product for management very well, so it is difficult to determine just how useful this product would be. If they are indeed able to identify clock gene SNPs that are linked to differences in run timing, then this would be a very exciting and potentially very useful tool in the study of phenotypic divergence in many different salmon populations. But without more information about the importance of distinguishing these runs for management in the San Francisco Bay populations of chinook, it is
-----------------	--

## Technical Review #2

	hard to know how useful this project would be for CalFED.
Rating	fair

## Additional Comments

Comments

## Capabilities

What is the track record of authors in terms of past performance? Is the project team qualified to efficiently and effectively implement the proposed project? Do they have available the infrastructure and other aspects of support necessary to accomplish the project?

Comments	The principal investigators' publication records suggest that they are qualified to carry out the proposed research. Both appear to have extensive experience with microsatellites, less so with SNPs. Based on the quality of the proposal, I have reservations about their ability to 'efficiently and effectively' implement the proposed research and turn it into a product that would be useful to CalFED. They likely have the necessary infrastructure at OSU.
Rating	fair

## Budget

Is the budget reasonable and adequate for the work proposed?

Comments	As the authors mention in the proposal, this kind of work is expensive. Given that, the proposed budget seems reasonable and well justified.
Rating	good

## Technical Review #2

### Overall

Provide a brief explanation of your summary rating.

<b>Comments</b>	The proposal is disorganized, incomplete, poorly written, and full of spelling and grammatical errors. In fact it is so bad that I cannot help but wonder if they accidentally submitted an early rough draft instead of the final one. If this is the case, perhaps they should resubmit it. If not, and the quality of the proposal reflects the quality of the products that will be submitted to CalFED, then I would not fund this proposal.
<b>Rating</b>	poor

# Technical Review #3

proposal title: SNPping the Cost and Proportional–bias error for Genetic Identification of Chinook Salmon Life Histories from Upstream Rivers to the San Francisco Bay Estuary

## Review Form

### Goals

Are the goals, objectives and hypotheses clearly stated and internally consistent? Is the idea timely and important?

Comments	<p>The goals, objectives, and hypotheses are very clearly stated. This proposal intends to add to the arsenal of microsatellite loci that are available for genotyping and population assignment of salmonids by characterizing nucleotide polymorphisms within a series of genes that are associated with the seasonal reproductive timing phenotype. It has long been known that timing of reproduction is a heritable phenotype in salmonids, and a number of microsatellites are known to be linked to one or more of these quantitative trait loci in rainbow trout. Now it appears that more detailed genomic information is available, and Dr. Banks and Dr. Bucklin intend to use these genes as non-neutral representative markers of the reproductive timing/run phenotype so that lineages which have been recognized in the past as distinct can continue to be managed as distinct lineages. The part of this proposal that is most interesting is that there is a growing acceptance, I hope, for the use of non-neutral markers in conservation genetics; if we know that a lineage has responded differentially to some selective pressure, or that there is a marker that unambiguously identifies a particular phenotype, there is no reason not to use this to better understand the evolution of that phenotype.</p>
Rating	excellent

### Technical Review #3

## Justification

Is the study justified relative to existing knowledge? Is a conceptual model clearly stated in the proposal and does it explain the underlying basis for the proposed work? Is the selection of research, pilot or demonstration project, or a full-scale implementation project justified?

Comments	Yes, the primary rationale for improving the genotyping ability is that mis-characterization of fish may lead to proportionally large errors when seasonal runs of very different sizes are compared. The hope is that not only will the probability of correct assignment increase with additional markers, but that their tie to genetic loci associated with spawning/run time will lead to diagnostic or near-diagnostic markers and assays. The basis for the proposed work is clear, and the selection of research populations is justifiable.
Rating	excellent

## Approach

Is the approach well designed and appropriate for meeting the objectives of the project? Is the approach feasible? Are results likely to add to the base of knowledge? Is the project likely to generate novel information, methodology, or approaches? Will the information ultimately be useful to decision makers?

Comments	They will certainly be able to meet the objectives of this proposal, it is feasible, but it appears that the SNP information is already available and this is primarily at the optimization stage for implementation - which loci work best for the least amount of money? If this set of markers is broadly applicable, that will be very useful for conservation of salmonid species; no information is given regarding the applicability of these markers to other salmonids, and perhaps the most cost-effective thing to do would be to establish the set of markers that would work
----------	---

### Technical Review #3

	in other species as well; it is not clear how much transspecific polymorphism there is at these loci from this proposal, so that may not be feasible.
Rating	very good

## Feasibility

Is the approach fully documented and technically feasible? What is the likelihood of success?  
Is the scale of the project consistent with the objectives and within the grasp of authors?

Comments	<p>The approach is not fully documented, in fact this proposal is in many places not completely written or edited. While I am not doubting the capability of the proposers, and it is technically feasible with a high likelihood of success, there is much information NOT given about how these SNPs are identified, screened, etc. from the sequence data that is available. What genetic data is presented is often un-rooted and poorly presented, and the very basic run phenograms that are presented for their SNP data, mitochondrial DNA data, MHC sequence data, microsats, allozymes etc. either completely ignore intraspecific variation in the actual sampling, or do not express how much intraspecific variation is within each group. It is very difficult to evaluate the "deeper branch lengths" when there are only 4 branches; more appropriate would be to express this as corrected distances (it is not clear if this is what is intended) or to evaluate the probability of co-assignment for all individuals as with some Bayesian assignment methods. I know that one of the proposers is an expert in statistical genotypic assignment tests, but I think the proposal itself leaves a lot to be desired. I know enough about genetics and these tests to understand what is probably intended, but I would not be surprised if another reviewer is puzzled by the gaps left here. It does certainly seem that success is likely since they are already funded to resolve SNPs for Chinook salmon, which leaves open the question of why so much</p>
----------	---

### Technical Review #3

	additional funding is requested.
Rating	good

## Monitoring

If applicable, is monitoring appropriately designed (pre–post comparisons; treatment–control comparisons)? Are there plans to interpret monitoring data or otherwise develop information?

Comments	I think once the loci are identified, the proper approach is being taken to ensure the best set of loci will be used for system-wide run ID usage.
Rating	excellent

## Products

Are products of value likely from the project? Are contributions to larger data management systems relevant and considered? Are interpretive (or interpretable) outcomes likely from the project?

Comments	It is unclear how applicable this data will be beyond Chinook salmon; while a valuable approach in and of itself, I think other species will probably have to go back to the 'clock gene' sequence data and reinvent the wheel with SNPs. In some cases, microsatellite loci that do crossamplify in salmonid species are known to be linked to QTL loci responsible for reproductive timing, and a full comparative analyses should probably be done at some point to determine which is the most cost-effective program for funding to be applied to. The overall product FOR Chinook salmon proposed here is, however, state-of-the-art and should dramatically improve assignment of individuals to distinct runs. More information on the heritability of these traits will be available by comparing linked vs. unlinked loci to see what genealogical patterns underlie these phenotypes. It is still unclear to me (from this proposal, anyway) whether there is an
----------	---

### Technical Review #3

	environmental determinant combined with historical isolation of lineages that has led to the different run times, or if it is an inborn switch completely associated with these genes. The family group assessment proposed on page 13 will help answer these questions, leading to interesting answers for evolutionary biology in general as well as the conservation and management community.
Rating	excellent

### Additional Comments

Comments	<p>Again, the proposal was poorly written, syntactically wrong in places, and in at least two spots an "x" was left in place of the actual value that was supposed to be pasted in. We've all done this, but there are enough errors left into this proposal that it makes me question it, though the science appears to be valid and well-designed otherwise.</p> <p>On page 8 of the executive summary, for example, it concludes with what is probably a comment to themselves "Give additional comments, information, etc. here". Most of the sentences on this page are not complete.</p>
----------	--

### Capabilities

What is the track record of authors in terms of past performance? Is the project team qualified to efficiently and effectively implement the proposed project? Do they have available the infrastructure and other aspects of support necessary to accomplish the project?

Comments	Dr. Michael Banks has an excellent track record in publishing and management advances, in particular with the development of assignment test software and theoretical advances. They clearly have the infrastructure and expertise to carry out this project.
Rating	

### Technical Review #3

	very good
--	-----------

## Budget

Is the budget reasonable and adequate for the work proposed?

Comments	<p>Frankly, given that the genes are already characterized and they are at the stage of PCR and genotyping optimization, the budget seems excessive at first glance. The ultimate, optimized cost of genotyping individuals should be around 0.08 per locus (their estimate; ABI was talking over a year ago about methods to genotype 96 loci at once for even less per-locus expenditure); at 48 loci and 3000 individuals, that is only going to cost about \$12,000.</p> <p>While travel funds seem appropriate, and the expenditure for the ABI 3730xl is entirely justified - enabling not just the genotyping, which could be done off-site in a high throughput facility, but the rapid screening of PCR optimization attempts. This is typically the most laborious and time-consuming step in going from a screen of genomic products to successful genotyping, and having this machine in their facility will dramatically improve their odds on this project plus is an investment for future projects that will pay off in a high-throughput analysis of any populations being studied.</p> <p>However, the estimate of supplies and equipment still seems outrageous. I am a geneticist and would love to know how a budget of \$3000 per researcher per month is reached. Primer pairs, when labeled with a fluorescent dye, cost \$100 each; polymerase, once considered very expensive, can be purchased for only about 0.20 per unit (the amount needed, nominally, to run a PCR reaction; frequently we can pull this off in half-reactions too). The other reagents, since the genome has already been mostly characterized for the loci they are discussing, are relatively cheap. Little</p>
----------	--

### Technical Review #3

	if any cloning would need to be done, just assessment of polymorphism, reliability, and scoreability. In other words, lots of man-hours and very tedious work, but could be done for far less funding in my opinion. This is an incredibly valuable task that is being proposed, but we have limited funds for conservation and management and unless this is going to be more applicable to broader taxonomic use, I would think the money could be spread out a little thinner.
Rating	fair

## Overall

Provide a brief explanation of your summary rating.

Comments	I'm torn. The science proposed could be valuable, evolutionarily interesting, and mark a real breakthrough in the genotypic assignment of salmonid seasonal runs. However, it is not clear from this proposal that the data is not essentially in hand, and so a much more modest funding could pay for the assessment of population status. I would recommend funding at a lower rate.
Rating	very good